I claim:

1. An apparatus containing a data structure representing a presentation, the data structure comprising:

a first audio channel representing an audio portion of the presentation after time scaling by a first time scale factor; and

a second audio channel representing the audio portion after time scaling by a second time scale factor that differs from the first time scale factor.

2. The apparatus of claim 1, wherein:

the first audio channel comprises plurality of frames;

the second audio channel comprises plurality of frames that are in one-to-one correspondence with the plurality of frames in the first audio channel; and

corresponding frames in the first and second audio channels represent the same time interval of the presentation.

3. The apparatus of claim 2, wherein each frame in the first audio channel is separately compressed using a first compression method.

4. The apparatus of claim 3, wherein the data structure further comprises a third audio channel representing the audio presentation after time scaling by the first time scale factor, wherein each frame in the third audio channel is separately compressed using a second compression method.

5. The apparatus of claim 1, wherein the data structure further comprises a data channel identifying graphics associated with the audio presentation.

6. The apparatus of claim 1, wherein:

the first audio channel comprises plurality of frames, each frame having an index value that identifies a time interval of the audio portion that the frame represents;

the second audio channel comprises plurality of frames, each frame in the second

channel having an index value that identifies a time interval of the audio portion that the frame represents.

7. The apparatus of claim 6, wherein each frame in the first and second data channels is separately compressed.

8. The apparatus of claim 6, wherein the data structure further comprises a data channel corresponding to a plurality of bookmarks, wherein each bookmark has index value and identifies graphics, the index value indicating a display time for the graphics relative to playing of the frames of the first or second audio channel.

9. The apparatus of claim 1, wherein the apparatus comprises a server connected to a network.

10. The apparatus of claim 1, wherein the apparatus comprises:

data storage in which the data structure is stored;

a decoder connected to receive a data stream from the data storage, the decoder converting the data stream for perceivable presentation; and

selection logic coupled to the data storage and capable of selecting a source channel for the data stream from among a set of channels including the first audio channel and the second audio channel.

11. The apparatus of claim 10, wherein the apparatus is a standalone device that operates on battery power.

12. An apparatus containing a data structure representing an audio presentation, the data structure comprising a plurality of audio channels representing the audio presentation after time scaling, wherein:

each audio channel has a corresponding time scale factor and includes a plurality of audio frames; and

each audio frame has a frame index that uniquely distinguishes the audio frame from

other audio frames in the same channel and identifies the audio frame as corresponding to specific audio frames in other audio channels.

13. The apparatus of claim 12, wherein audio frames that are in different channels and have the same frame index represent the same portion of the audio presentation.

14. A method for encoding audio data, comprising:

performing a plurality of time scaling processes on the audio data to generate a plurality of time-scaled audio data sets, each time-scaled audio data set having a different time scale factor; and

generating a data structure containing a plurality of audio channels respectively corresponding to the plurality of time scaling processes, wherein content of each of the audio channels is derived from the time-scaled audio data set resulting from performing the corresponding time scaling process on the audio data.

15. The method of claim 14, wherein generating the data structure comprises:

partitioning each time-scaled audio data set into a plurality of frames;

separately compressing each frame to produce compressed frames; and

collecting the compressed frames into the plurality of audio channels, each audio channel having a corresponding one of the different time scale factors.

16. The method of claim 15, wherein all frames resulting from the partitioning correspond to the same amount of time in the audio data.

17. The method of claim 15, wherein separately compressing each frame comprises applying a plurality of different compression processes to generate a plurality of compressed frames from each frame.

18. The method of claim 17, wherein collecting the compressed frames produces audio channels such that in each audio channel, all compressed frames in the audio channel have the same time scale and compression process.

19. A method for playing a presentation, comprising:

loading a first frame from a source into a player via a network, the first frame representing a first portion of the presentation after scaling by a first time-scaling factor, wherein the first audio frame has a first channel index value that identifies the first audio frame as being scaled by the first time scaling factor;

playing the first portion of the presentation based on data from the first audio frame;

receiving a request to change playing from the first time scaling factor to a second time scaling factor;

requesting from the source a second audio frame that has a second channel index value that identifies the second frame as being scaled by the second time-scaling factor; and

playing the second frame after the first to provide a real-time change in the time-scale of the presentation.

20. The method of claim 19, wherein the first frame has a first frame index value that identifies the first portion of the presentation that the first audio frame represents, and the second frame has a second index value that identifies a second portion of the presentation that the first audio frame represents.

21. The method of claim 20, wherein the second index value immediately follows the first time index value

22. The method of claim 19, wherein channel index values of frames further indicate respective compression processes for the frames, and wherein the method further comprises:

determining available bandwidth on the network; and

selecting the second channel index value from a plurality of channel index values that identify the second time scaling factor, wherein the second channel index indicates a compression process provides highest audio quality at the available bandwidth.

23. The method of claim 19, wherein channel index values of frames further indicate respective compression processes for the frames, and wherein the method further comprises:

determining available bandwidth on the network;

selecting a third channel index value from a plurality of channel index values that identify the second time scaling factor, wherein the third channel index indicates a compression process provides highest audio quality at the available bandwidth;

requesting from the source a third audio frame that has the third channel index value, which identifies the third audio frame as being time-scaled by the second time-scaling factor; and

playing the third frame after the second frame to provide a real-time change in the time-scale of the presentation

24. A method for playing an audio presentation on a receiver that is connected via a network to a source having a multi-channel data structure representing the audio presentation, the method comprising:

determining available bandwidth on the network;

selecting a first channel of the multi-channel data structure from a plurality of channels that represent the audio presentation after time-scaling by a desired time-scaling factor, wherein the first channel contains data that is compressed using a compression process that provides highest audio quality at the available bandwidth;

receiving a first frame from the first channel; and

playing the first frame.

25. The method of claim 24, further comprising:

determining bandwidth available on the network after receiving the first frame;

selecting a second channel of the multi-channel data structure from the plurality of channels that represent the audio presentation after time-scaling by the desired time-scaling factor, wherein the second channel contains data that is compressed using a second compression process that provides highest audio quality at the bandwidth available after receiving the first frame;

receiving a second frame from the second channel; and

playing the second frame after playing the first frame.

26. A method for controlling display of web pages, comprising:

assigning a series of web pages to respective index values of audio data that represent an audio portion of a presentation;

playing audio generated from the audio data; and

displaying each web page in response to the playing reaching in the audio data an index value assigned to the web page.

27. The method of claim 26, wherein assigning the series of web pages comprises:

partitioning the audio data into a series of frames;

assigning a different index value to each of the frames; and

assigning each web page to the index value of a frame, wherein the web page is to be displayed while the frame is played.

28. The method of claim 26, wherein assigning the series of web pages comprises creating a data structure including:

an audio channel containing audio frames that together constitute the audio data; and

a data channel containing for each web page, a link to the web page and frame index value identifying an audio frame corresponding to the web page.

29. The method of claim 26, wherein assigning the series of web pages to respective index values comprising assigning each web page to a start index value and a stop index value, wherein the web page is to be displayed during playing of frames having index values between the start index value and the stop index value.

30. A method for authoring a presentation for playback on a computing system, comprising:

assigning time index values to audio data for the presentation;

assigning a range of the time index values to each image represented by graphics data for the presentation; and

constructing a file containing the audio data and the graphics data, wherein the file has a format indicating display of each image occurs during playing of the audio data that has

assigned time index values in the range assigned to the image.

31. The method of claim 30, wherein the graphics data comprises a link that identifies data available on a network, and display of the image associated with the link comprises retrieving data that the link identifies.

32. The method of claim 31, wherein the link identifies a web page, and display of the image associated with the link further comprises displaying the web page.

33. The method of claim 30, wherein the graphics data comprises image data that is embedded in the file, and displaying the image comprises displaying an image that the image data represents.

34. The method of claim 30, wherein:

assigning time index values to the audio portion comprises partitioning the audio data into a plurality of frames, wherein each frame has a time index value according to an order for playing of the frames; and

constructing the file comprises collecting the frames into an audio channel.

35. The method of claim 34, further comprising collecting the graphic data in a data channel.

36. The method of claim 30, wherein assigning the ranges of the time index values to the images comprises:

representing a time span of the audio data;

selecting a point in the time span; and

selecting one of the images to be assigned to the point selected.